The Evolution of "Kantian Trait": Inferring from the Dictator Game

Lorenzo Cerda Planas

The Evolution of a "Kantian Trait": Inferring from the Dictator Game

Lorenzo CERDA PLANAS

2015.32

ISSN : 1955-611X

The Evolution of a "Kantian Trait": Inferring from the Dictator Game

Lorenzo Cerda Planas*

Last revision: March 18^{th}, 2015

Abstract

The aim of this paper is twofold. Starting from the population dynamics literature, which usually finds the resulting distribution of a trait in a population, according to some parents' preferences, I answer the inverted question: Which preference function would yield into a given trait distribution? I solve this using a continuous trait, instead of finite types of agents. Using this result, I connect this transmission theory of social traits with the well-known results of Dictator Game (DG) experiments. I use a specific definition of a Kantian trait applied to DG results, and determine the distribution of this trait that is commonly found in these experiments. With these two ingredients, I show that homo-oeconomicus parents have a greater 'dislike' or disutility of having offspring with different traits from them compared to their Kantian counterparts. This could be a result of myopic empathy being stronger in homo-oeconomicus parents, driving this dislike of difference.

JEL Classification: C62, C63, C73, C61, D64.
Keywords: Population dynamics, Kantian morale, evolutionary equilibrium.

*Paris School of Economics, C.E.S. - Université de Paris 1 Panthéon-Sorbonne; 106-112, Boulevard de l'Hôpital, 75647 Paris Cedex 13, France. Email: lorenzo.cerda@gmail.com

1 Introduction

The goal of this paper is twofold. First, using the transmission theory of social traits, specifically the population dynamics tool-kit, I develop the transmission dynamics of a trait that lies in a continuous segment. This trait is transmitted using some myopic preference of parents (concerning their offspring utility) as in Bisin & Verdier (2000) [1]. Within this set-up, the literature usually develops to find the equilibrium. I invert this question and ask: Which parents' preference function would lead to a given population distribution? I solve the conditions that have to be met and I also develop an algorithm to solve for more complex cases. On the other hand, and using the previous results, I connect the transmission theory of social traits with the results of a well known experiment in economics, the Dictator Game (DG). In order to do so, I use the experimental results of the DG in order to infer the distribution of a moral trait in a society. In this case, the moral trait is what I call a Kantian morale. It is possible to map the responses of DG experiments into these types of moral traits. Assuming that distributions of actual societies are the result of a long evolutionary transmission process of these kind of traits, and using the aforementioned results, I ask if there is such an evolutionary process that could explain the results of DG experiments. It turns out that such a process could exist (mathematically speaking), and if this is the case, it would imply that homo-oeconomicus parents have stronger feelings about having offspring similar to them, than their Kantian counterparts. This result can turn out to be important when dealing with environmental challenges, where individual provision of environmental public goods and coordination are greatly needed in order to sustain a clean environment.

The literature on social trait transmission uses the population dynamics tool kit in order to model how the next generation will be, according to the present state of the distribution of traits and the 'forces' involved in the evolving process. For example, Bisin & Verdier (2000) [1] suppose that agents of the present generation, using a myopic empathy when considering their offspring's utility, try to transmit their own traits to their children.(1) This myopic empathy means that parents evaluate their offspring's utility according to their own. This implies that parents will evaluate their children's actions using their own utility functions, which in turn will yield to a lower utility of their offspring, from the parents' point of view, if their traits differ from those of their parents. I use this starting point and I make two modifications: I assume there exists a 'myopic dislike' function $v(\cdot)$ of the agent towards his or her child's trait when this trait is different from their own; and I assume that this trait can be modelled with one variable positioned in a continuous line between zero and one. This departs from the literature, where usually

(1)This type of trait transmission has also been used in Bisin & Verdier (2001) [2], Hauk & Saez-Marti (2002) [7] and Saez-Marti & Zenou (2012) [11].

authors study a specific scenario with a finite number of types of agents, usually being equal to two or three.

On the other hand, the Dictator Game (DG) is one of the simplest and most replicated economic experiments in the game theory area. In a nutshell, the experiment consists of recruiting two people and giving an amount of money (or another valuable thing) to one of them, chosen randomly. The person who received the money is called the Proposer (or Dictator) and he is asked to share some fraction (or none) of this money with the second person, called the Responder. After this, the money is split according to this decision. Therefore, the Responder has no say in this game; we are only interested in the Proposer's decision. According to the homo-oeconomicus theory, the Dictator should never give a dime, but these experiments show a constant and considerable amount of people sharing some part, even to a 50/50 proportion or more. Different explanations of this phenomenon exist. Assuming that the Proposer has no direct or indirect relationship with the Responder, which is usually the experiment set-up, these explanations point to the idea of a social norm and/or a moral trait, both of which are transmitted between generations. Following the idea of Cerda (2015) [3] regarding a Kantian morale, where agents are endowed with a Kantian trait that lies in a continuous spectrum, I can map the DG responses to a Kantian trait level. In this framework, I define a Kantian person as an agent maximizing his utility *assuming* that everyone acts as he does, in contrast with the homo-oeconomicus agent, who just maximizes his utility in a selfish manner. An interesting result of the DG experiments is that there is a kind of polarization of the distribution, having roughly one third of the population acting in a purely homo-oeconomicus way (giving nothing), another third or so sharing half of the pie (or even more), which I translate to being fully Kantian, and the rest of the people distributed, almost uniformly, between these two extremes.

With these two ingredients in mind, the objective of this paper is to rationalize the evolution of the distribution of the Kantian trait to one that could account for the results of DG experiments. In order to do so, I will first develop a continuous model of population dynamics, using the discrete one as a starting point. I assume that there exists a $v(\cdot)$ function that represents the parent's 'dislike' or loss in utility of having a child with a trait different from their own. The input of the function is the difference of the child's trait compared to his or her parent's trait. A positive amount means that the child is more Kantian than his or her parent, where a negative one means the inverse case. The $v(\cdot)$ function does not need to be symmetrical in zero. Actually, we will see that in order to match the DG experiment's results, it will not be symmetrical. With this, I find equilibrium conditions of the dynamics where the distribution of the population stops evolving. I am assuming here that the results of the DG experiments are actually in equilibrium, or quite close to it. With this set-up, I show some simple results concerning possible solutions of $v(\cdot)$ functions, which evolve into a given equilibrium distribution of the population. This distribution lies in the continuous segment $[0,1]$, and it can have points of high concentration, modelled with Dirac deltas. The only condition for the distribution is that its integral has to be equal to one, as in any distribution.

Having the basic results of the equilibrium conditions and some properties of $v(\cdot)$, I move into finding $v(\cdot)$ function(s) that can make a population distribution evolve into what we have as results in DG experiments. It turns out that solving this problem analytically is not plausible for this asymmetric case, and I have to rely on simulations. Here I develop an algorithm that takes the resulting (final) distribution as input and an initial guess for $v(\cdot)$; using an iterative process, it converges to a $v(\cdot)$ function that meets the equilibrium conditions. The paper starts in Section 2 with the basic discrete model and explains how I transform it to its continuous counterpart. This section also includes some basic properties of the solutions. In Section 3, I move to the general case, introducing the algorithm and showing the solution for the case of the DG experiments' results. Section 4 concludes.

2 The Initial Model

2.1 The starting point

I begin with the trait transmission, using as a starting point the model used by Bisin & Verdier (2000) [1]. They used what it is called 'vertical' and 'oblique' cultural transmission. The idea is that parents will try to transmit their own trait to their children, which will be effectively transmitted with some probability τ. If they fail to transmit the trait, then with probability $(1-\tau)$ the child is matched randomly with an individual of an old generation and adopts his trait. In their paper, the authors deal with a population consisting of only two types of people, therefore having the vertical transmission probabilities of τ^a and τ^b (one for each type). By calling q_t the share of population of type a in period t (and therefore having $(1-q_t)$ the share of population of type b), the transition probabilities P_t^{ij} of type i having a child of type j, are easily calculated (as in [1]):

$$P_t^{aa} = \tau^a + (1-\tau^a)q_t \qquad P_t^{ab} = (1-\tau^a)(1-q_t) \tag{2.1}$$
$$P_t^{bb} = (1-\tau^b)(1-q_t) \qquad P_t^{ba} = (1-\tau^b)q_t \tag{2.2}$$

Given these probabilities, the share of type a in period $t+1$ is derived too:

$$q_{t+1} = q_t + q_t(1-q_t)[\tau^a - \tau^b] \tag{2.3}$$

This is the standard evolution equation found in population dynamics, as in for example Sigmund (1986) [12], Silverberg (1997) [13], Hofbauer & Sigmund (2003) [8], and Harper (2009) [6]. Still following Bisin & Verdier's paper, the parent will bear a cost when socializing their child with a given trait. In this case, this cost is denoted by $H(\tau^i)$, depending on the socialization effort τ^i. The parent chooses τ^i that maximize

$$\beta[P_t^{ii}V^{ii} + P_t^{ij}V^{ij}] - H(\tau^i) \tag{2.4}$$

where β is the discount rate and V^{ij} is the utility of a child of type j perceived by a parent of type i. Again following the literature, I assume that parents act according to 'imperfect empathy', meaning that they evaluate their child's utility through their own imperfect lenses. This is the point where I depart from the literature. First, I assume that V^{ij} is constant in time and well-known by the parents of type i. These values do not depend on the composition of the society, as they do in some cases in the literature. In any case, the maximization in Eqn. 2.4 does depend on the society composition, through the transition probabilities P_t^{ij}. Secondly, I 'normalize' the value of V^{ij} such that instead of being the child j's utility viewed in parent i's eyes, it will be the *difference* of this aforementioned utility and the parent's. In other words, I define $V(i,j) = V^{ij} - V^{ii}$. Therefore $V(i,j)$ is the loss of utility that a parent of type i has, when having a child of type j. This will turn out to be handy in the generalization to n types of agents.(2)

Returning to the population dynamics and observing Eqn. (2.3), it is easy to see that if both types coexist ($q \neq 0 \wedge q \neq 1$), then the system stops evolving when $\tau^a = \tau^b$, meaning that both types of parent are exerting the same amount of effort in socializing their children.

2.2 Extending the model

Now I will add more types to the model, then transform it into one with types of people lying in a continuous segment. First, we have n types of agents; let us have them ordered, as for example with the natural numbers. In other words, we have different degrees of a trait and the n types just signify the strength of this trait. As explained in the Introduction, I will relate this type i with how Kantian or homo-oeconomicus a person is, as in Cerda (2015) [3]. Hence, we can think of having n types $1, 2, \ldots, n$, where $i = 1$ means a fully homo-oeconomicus person and $i = n$ means a fully Kantian one. Those in between are ordered in the sense that if $j > i$, it means that type j is more Kantian than i,

(2) This modification does not change the maximization problem. It is easy to verify this by replacing the term P_t^{ii} with $1 - P_t^{ij}$ in Eqn. (2.4). Since V^{ii} is constant with respect to τ^i, we have the same maximization program. In the general case, we make $P_t^{ii} = \sum_{j=1}^{n}(1 - P_t^{ij})$.

although this idea can be applied to any trait that allows an ordering.

I also assume that $V(i,j)$ is a 'dislike' function, as previously mentioned, and that it increases with the difference of i and j. This means that the more different a child turns out to be (with respect to his parent), the bigger the disutility that his parent bears. We also have, following its definition, that $V(i,i) = 0$, $\forall i$.

Following the previous notation, let q_t^i be the share of type i in the population at time t. It is now easy to extend the two type equations (Eqns. (2.1) and (2.2)) into n types. We can now write the transition probabilities:

$$P^{ii} = \tau_i + (1-\tau_i)q_t^i \qquad P^{ij} = (1-\tau_i)q_t^j \qquad \forall i \neq j \tag{2.5}$$

We can also construct a matrix V with its elements being: $V^{ij} = V(i,j)$ and the effort vector $\vec{\tau} = (\tau_1, \tau_2, \dots, \tau_n)^T$. The maximization problem for each agent j will be:

$$\max_{\tau_j} \beta \Big(\sum_{i=1}^{n} P^{ji} \cdot V^{ji} \Big) - H(\tau_j) \tag{2.6}$$

Noting that the diagonal of V is full of zeros (since $V^{ii} = 0$), we can redefine the matrix P without changing the system of equations by having $P = (\vec{1} - \vec{\tau}) \cdot \vec{q}^T$, where $\vec{1}$ is a vector of ones and $\vec{q}$ is the vector composed by the shares of each type (and hence, the sum of its elements equals one). With this, the solution of the maximization problem is:

$$\beta \frac{\partial((P \cdot V^T)_j)}{\partial \tau_j} = -\beta \sum_{i \neq j} q_t^i \cdot V(j,i) = H'(\tau_j) \tag{2.7}$$

For simplicity, let me assume that $H(\tau) = 1/2\tau^2$, and since $V(j,i)$ is non-positive, that $V(j,i) = -c \cdot v(j,i)$, with $c > 0$ constant and $v(j,i) \geq 0$. With all these assumptions, we can write the solution of the maximization problem:

$$\tau_j = \beta \cdot c \sum_{i \neq j} q_t^i \cdot v(j,i) \tag{2.8}$$

We can now return to the dynamic of the population. It is easy to show that the general solution the evolution of q_t^j is:(3)

$$\Delta q_t^j = q_{t+1}^j - q_t^j = (\tau_j - \overline{\tau}) \cdot q_t^j \qquad \text{with} \qquad \overline{\tau} = \sum_{i=1}^{n} \tau^i q_t^i \tag{2.9}$$

(3) For details, please see Appendix A.

It is worth noting that any constant factor that shows up in Eqn. (2.8) will also appear in $\overline{\tau}$ and therefore in Δq_t^j. This will only change the speed of the evolution process; the equilibrium will be the same. Following this, we do not need to worry about the constant factors β and c in Eqn. (2.8). From Eqn. (2.9) it is straightforward that at equilibrium, when $\Delta q_t^j = 0$, for all surviving types j (i.e. $q_j^* > 0$), $\tau_j = \overline{\tau}\ \forall j$. In words this means that the transmission effort τ_j that, for those types that did not disappear in the evolution process ($q_j^* > 0$), is the same (or constant). This is the equilibrium condition.

The continuous case

The idea is to extend the model to one with continuous types of agents. Now agents' type α will lie in the continuous segment $[0,1]$. Therefore, the distribution of the population will be $f(\alpha) \geq 0$, with

$$\int_0^1 f(\alpha)d\alpha = 1 \tag{2.10}$$

The condition that $\tau^k = \overline{\tau}$ for all types of agents can be rewritten as following, using the result of Eqn. (2.8):

$$\int_0^1 v(j,\alpha)f(\alpha)d\alpha = \tau(j) = \text{constant} \qquad \forall j \mid f(j) > 0 \tag{2.11}$$

In other words, this means that a population whose dynamic is 'defined' by the function $v(\cdot)$ will converge into a distribution $f(\alpha)$ when condition (2.11) is met. As with other results in population dynamics, there may exist more than one converging distribution $f(\alpha)$ for a given dynamic, and the final distribution of the population depends on the initial state of the population distribution (see for example Zeeman (1980) [14] and Friedman (1998) [5]). Another way to use condition (2.11) is to ask: Which $v(\cdot)$ dynamics (if there are any) would converge to a population defined by $f(\alpha)$?

In order to start answering this question, let us try two simple examples. First, let us answer this question with the following $f(\alpha)$:

$$f(\alpha) = \frac{1}{2}\delta(0) + \frac{1}{2}\delta(1) \tag{2.12}$$

where $\delta(x)$ is Dirac delta function[(4)] centred in x. This means that $f(\alpha)$ is a polarized population in which half of the people are concentrated in $\alpha = 0$ and the other half in $\alpha = 1$. In order to simplify the formulation, note that $v(j,\alpha)$ is actually a function of the difference of j and α. Hence, we can rename it $v(j-\alpha)$, which will be handy later on. It turns out

[(4)] Dirac delta function or δ function is zero everywhere except at zero, with an integral of one over the entire real line.

that condition (2.11) simply becomes $v(j) + v(1-j) = constant$, for $j = 0$ and $j = 1$. This is just because $v(0) + v(1) = v(1) + v(0)$, which is always true, no matter which function $v(\cdot)$ we use. This means that this polarized distribution is a solution of **any** $v(\cdot)$ dynamics.

Let us now try with a more complex example:

$$f(\alpha) = C_1\,\delta(0) + C_2 + C_1\,\delta(1) \qquad \text{with } \; 2C_1 + C_2 = 1 \tag{2.13}$$

In this one we have also a (semi)polarized distribution, where there are people in the intermediate values of α. These intermediate people are distributed uniformly. Solving again for condition (2.11)[5], we get the following:

$$C_1(v(j) + v(1-j)) + C_2\Big(\int_0^j v(j-x)dx + \int_j^1 v(x-j)dx\Big) = \text{constant} \tag{2.14}$$

Functional solutions for this integral equation can be hard to find. A way to find (at least) some of them would be to transform it into a differential equation and try to find some solutions in that domain. By calling $w(j) = \int_0^j v(x)dx$, and hence $w'(j) = v(j)$, we can transform the previous equation into:

$$C_1 \cdot \big(w'(j) + w'(1-j)\big) + C_2 \cdot \big(w(j) + w(1-j)\big) = \text{constant} \tag{2.15}$$

Since this particular distribution $f(\alpha)$ is symmetric, I focused on searching for solutions that are symmetric around zero, meaning that $v(-j) = v(j)$. Two solutions for this equation are:[6]

$$v(j) = K - a \cdot c \cdot e^{-a|j|} \qquad \text{with} \quad a = C_2/C_1 \;\wedge\; K = a \cdot c \tag{2.16}$$

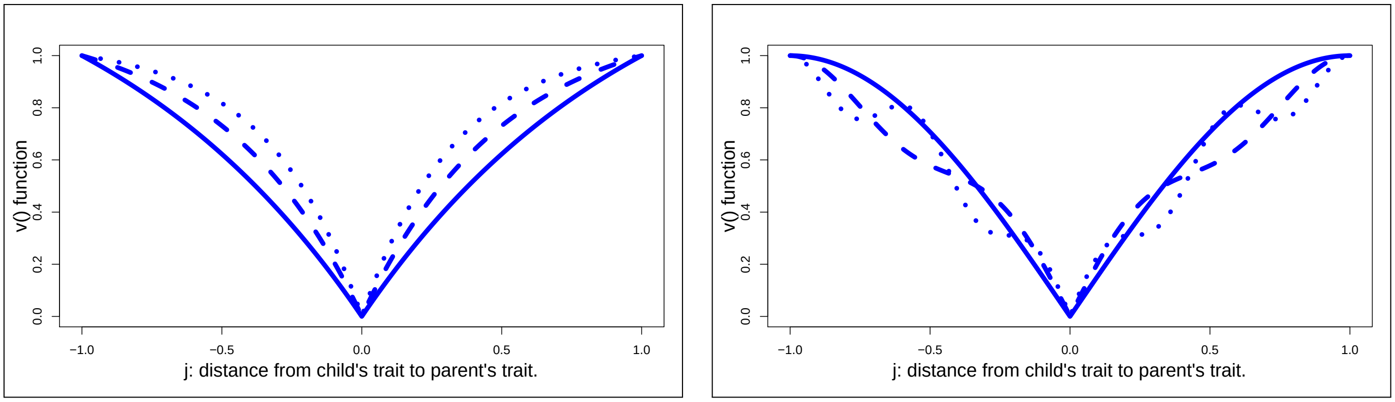
$$v(j) = b \cdot sin\Big(\frac{(4k-3)\pi}{2} \cdot |j|\Big) \qquad \text{with} \quad b > 0 \text{ (see footnote (6))} \;\wedge\; k \in \mathbb{N} \tag{2.17}$$

Constant K in (2.16) is such that $v(0) = 0$, and c is an arbitrary positive number. Taking these solutions into account, we now have a family of solutions, just by making a positive linear composition of different cases of (2.16) and (2.17). It is worth noting that in the case of the sine solution (2.17), when adding a higher-order sine (higher values of k), one should pay attention to its factor, since we want the final $v(j)$ to be an increasing function. In Fig. 1, some examples for both cases are plotted.

---

(5) For details, please see Appendix B.

(6) The parameter b in solution (2.17) is linked to C_1, C_2 and k, and therefore, this is a solution for a specific combination of C_1 and C_2. In order to have a generic solution for any C_1 and C_2, a linear combination of sines is needed (with different values of k). For details on this point and on the development of the solutions, please refer to Appendix B.



(a) Examples of solution in (2.16).

(b) Examples of solution in (2.17).

Figure 1: Some examples of solutions for $v(j)$.

3 General Solution

It is easy to see that finding an algebraic solution for any given distribution $f(\alpha)$ is not possible, and therefore I will rely on simulations and a numerical solution. In order to do so, I propose an algorithm that finds a solution, for a given distribution $f(\alpha)$. The algorithm starts from an initial guess for $v_0(j)$ and converges to a solution of $v(j)$. As expected, the solution found will depend on the initial guess.

The algorithm is the following:(7) take an initial guess of $v(j) = v_0(j)$. Compute $\tau_0(j) = \int_0^1 v_0(j-\alpha)f(\alpha)d\alpha$ for $j \in [0,1]$ (which is the expression in (2.11)). If $\tau_0(j)$ is constant(8) (for $0 \leq j \leq 1$), then $v_0(j)$ is a solution. If not, define an *adjustment function* as:

$$a_0(j) = \frac{1}{\tau_0(1-2j)} \tag{3.1}$$

and compute a new guess function $v_1(j) = a_0(j) \cdot v_0(j)$. Compute the new $\tau_1(j)$ and restart the process.

The intuition for choosing this adjustment function is the following: We want to find a $v(j)$ function that, when plugged into the equilibrium condition (Eqn. (2.11)), gives us a constant value for $\tau(j)$. Now, when moving j between zero and one, this integral (the equilibrium condition) is the multiplication of $f(\alpha)$, and $v(j-\alpha)$ that 'moves' along it. Therefore, if we do not get a constant value for $\tau(j)$, we want to correct $v(j)$ such that

(7)Properties of the algorithm can be found in Appendix C. I show that if at some iteration k, $v_k(j)$ is a solution, the algorithm converges, and that if it converges, then $v_k(j)$ is a solution. I do not show that the algorithm will always converge, although simulations' results suggest it does. The algorithm was programmed in R language and is available upon request.

(8)Or close enough to be constant, depending on the desired precision.

it does. The multiplicative inverse accounts for this purpose. As to for the factor 2 in j, this is due to the fact that the range of $v(j)$ is between -1 and 1, where it is only $[0,1]$ for $\tau(j)$. Finally, the $(1-2j)$ term, which is just a vertical mirror of the function (expanded by 2, as just explained), comes from the fact that we are making this 'sweep' in the inverse direction (note α in $v(j-\alpha)$, inside the integral in condition (2.11)).

With this algorithm it is possible to find solutions for more complex distribution functions – in particular, cases where the distribution is not symmetric. Returning to the initial goal, we can find $v(j)$ functions that account for the evolution of a Kantian trait that can explain, at least in part, the behaviour found in the Dictator Game experiments. Different papers can be used as sources of information, and I focus on Engel (2011) [4], which is a meta study on Dictator Games. It is useful for my purpose since it compiles a large amount of experiments and responses. In Fig. 2 of this work, he shows the result of 328 treatments with full range information, composed of the answers of 20,813 people. I replicate this in Fig. 2a in the coming pages.

Recalling the Introduction, the DG is played between two agents, who will be chosen randomly to be the Proposer (Dictator) and Responder. We are interested in the Proposer's action,(9) and each agent has a 50% chance to be elected as such. He or she will choose to give a share of the money they receive when chosen Proposer. Therefore, in order to use this information, I will transform the DG responses (give rate) into what I call Kantian trait. As stated in the Introduction and following Cerda (2015) [3], I can associate a Kantian 'measurement' to each respondent based on his give rate. Therefore, each person is defined by a α value (his Kantian trait) when he is acting acting in a way that maximizes the following utility function:

$$U(\cdot) = (1-\alpha)\cdot U_H(\cdot) + \alpha \cdot U_K(\cdot) \tag{3.2}$$

where $U_H(\cdot)$ is the utility function for an homo-oeconomicus agent and $U_K(\cdot)$ is the one for the Kantian person. A Kantian person is defined as one who maximizes his utility *assuming* that everyone acts as he does. One useful feature of this approach is that it transforms different experiments' results into a single measurement of what we could call a Kantian trait. In the case of the DG, the utility function is a function that transforms money (or the asset used in the DG) into utility. Typically this is a consumption utility function, with the classic properties of being increasing and concave. Using different utility functions, one can map different give rates (between zero and one half) into Kantian traits α, which will lie in the segment $[0,1]$. To see this, let $u(\cdot)$ be the consumption utility function in the DG experiment, γ the fraction shared by the Proposer (give rate), and C

(9) The Responder has no say in this game.

the amount received by this agent. Therefore we have, for the homo-oeconomicus agent, the fully Kantian agent, and the general case, the following maximization programs:

$$\max_{0\leq\gamma\leq 1} \underbrace{\frac{1}{2}u((1-\gamma)C)}_{U_H(\cdot)} \quad \rightarrow \quad \gamma^* = 0 \tag{3.3}$$

$$\max_{0\leq\gamma\leq 1} \underbrace{\frac{1}{2}u((1-\gamma)C) + \frac{1}{2}u(\gamma C)}_{U_K(\cdot)} \quad \rightarrow \quad \gamma^* = 1/2 \tag{3.4}$$

$$\max_{0\leq\gamma\leq 1} (1-\alpha)\cdot\frac{1}{2}u((1-\gamma)C) + \alpha\cdot\big[\frac{1}{2}u((1-\gamma)C) + \frac{1}{2}u(\gamma C)\big] \tag{3.5}$$

The $1/2$ value comes from the chance of being chosen as Proposer or Responder. Therefore, for the homo-oeconomicus case, he or she only evaluates when they are chosen as Proposer. From the point of view of the Kantian person, who decides assuming that everyone will act as they do there are gains when being a Proposer (with one half chance) and Responder (the other half), hence the formulation in (3.4). For the first two cases we have two straight solutions, no matter which utility function $u(\cdot)$ we use. For the homo-oeconomicus agent ($\alpha = 0$), the give rate is zero ($\gamma^* = 0$), as in Eqn. (3.3). For the fully Kantian agent ($\alpha = 1$, Eqn. (3.4)), the give rate is equal to one half ($\gamma^* = 1/2$). As for those agents with $0 \leq \alpha \leq 1$, the value of γ^* that solves the maximization program, in Eqn. (3.5), is:

$$u'(\gamma C) = \alpha u'((1-\gamma)C) \tag{3.6}$$

Therefore, the relationship between α and the pair (γ, C) will depend on the choice of $u(\cdot)$, as in the following examples:

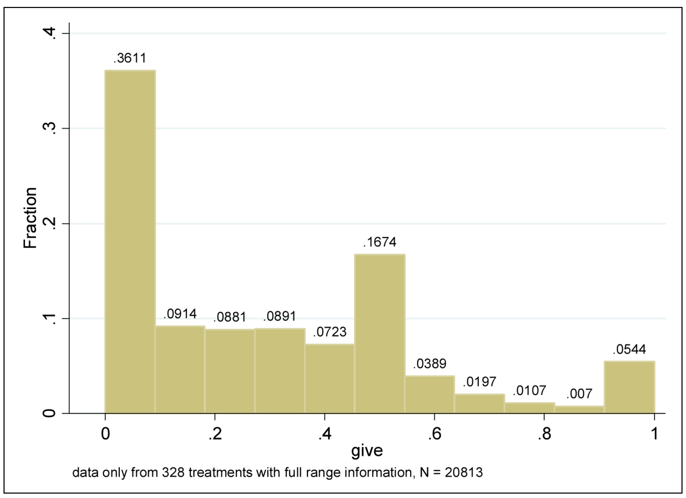
$$u(c) = ln(c) \quad \rightarrow \quad \alpha = \frac{\gamma}{1-\gamma} \tag{3.7}$$

$$u(c) = \frac{1}{1-\epsilon}c^{1-\epsilon} \quad \rightarrow \quad \alpha = \Big(\frac{\gamma}{1-\gamma}\Big)^{\epsilon} \tag{3.8}$$

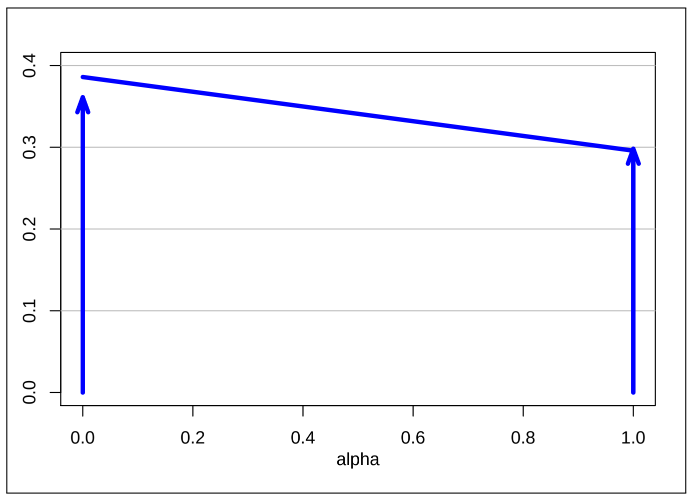
In these cases, I used a Constant Relative Risk Aversion (CRRA) utility function, with ϵ as the measure of risk aversion ($\epsilon = 1$ in the first case). Given this special form of the utility function, we find that the relationship does not depend on the amount of money to share, but only on give rate. In general, give rates do not substantially change with the amount to share, except when it comes to big ranges in the amount to divide, as in Novakova and Flegr (2013) [9].(10)

(10)Usually DGs are played with different amounts of money involved, although there is not normally a substantial difference among these values. The authors investigate a bigger difference, ranging approxi-

Concerning the value of ϵ, figures too different from $\epsilon = 1$ will produce a loss of information (when transforming from γ to α). With this consideration, and bearing in mind that the log function has been widely adopted in the literature, I use an ϵ equal to one.(11) Finally, for those people that gave more than one half, we should ask ourselves if we should discard them when transforming the distribution to a Kantian equivalent or, as I do, assume that these people are fully Kantian and the extra give comes from either miscalculation, misunderstanding, another motivation, or a combination of these factors. This assumption is in line with some detailed results found in O'Garra & Krantz (2014) [10],(12) as in the follow-up questions (which were open-ended) for dictators. These questions were aimed at identifying the reasons for their choices and are in line with my previous statement. In any case, if we were to discard this information, the main result does not change much (results are included in Appendix D). In Fig. 2 below the original results of Engel and its transformation to Kantian measurement are plotted.(13) The two



(a) Distribution of give rates, from Engel (2011).



(b) Distribution in Kantian measurement.

Figure 2: Results from DG meta study and its equivalent in Kantian measurement.

arrows in $\alpha = 0$ and $\alpha = 1$ in Fig. 2b are Dirac delta functions representing the two groups of people that cluster in these values, where the straight line is the density function of people with a Kantian measurement α in between. To arrive at this distribution, we have to find the parameters of this straight line (intercept and slope) that, when converted to their equivalent give rates γ, best fit the data. It turns out that a quasi-linear decreasing distri-

mately between \$1 and \$10,000, although their study is a survey and no real money was actually provided.

(11)For some examples and a deeper explanation, please see Appendix E.

(12)I greatly thank Tanya O'Garra of the Center for Research on Environmental Decisions, Earth Institute, Columbia University for providing me with the data from her Dictator Game study.

(13)I greatly thank Professor Dr. Christoph Engel of the Max Plank Institute for Research on Collective Goods, for providing me with the data of his meta study on Dictator Games.

bution fits the data quite well (which was suggested by the original histogram in Fig. 2a); this distribution does not vary much when we use different transformations between α and γ. The slope and intercept will change a bit, maintaining a decreasing trend. On the other hand, the cluster of people giving zero are translated into the Dirac delta in $\alpha = 0$, and people giving 50% or more are transformed into the Dirac delta in $\alpha = 1$. These two magnitudes are, of course, invariable with respect to the previous transformation. In the case of the depicted distribution of α (Fig. 2b), I used a simple linear relationship, although different transformations lead to very similar results.

Setting this distribution as $f(\alpha)$ and using the aforementioned algorithm, I derive some solutions for $v(\cdot)$, which are depicted in Fig. 3. There are some interesting things

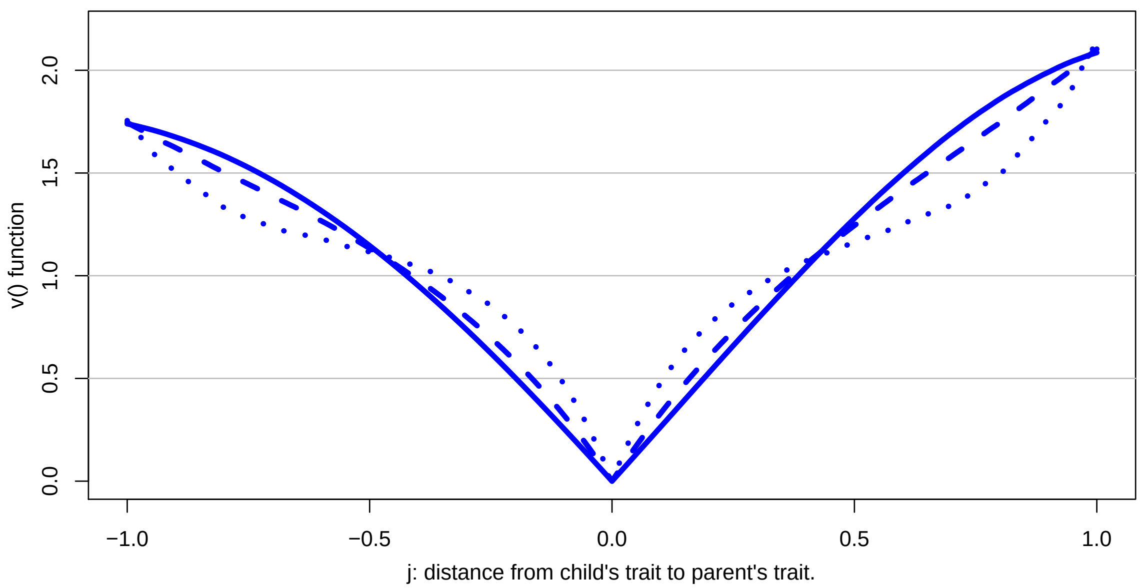


Figure 3: Examples of solutions for $v(\cdot)$.

to note. First (and somehow expected), the $v(\cdot)$ function is not symmetric in zero. This comes from the fact the $f(\alpha)$ is not symmetric either, and therefore the only way for $v(\cdot)$ to match condition 2.11 is not to be symmetric[14]. Also, even if we can have different functions $v(\cdot)$ that meet this condition, both sides of it have a similar shape with respect to zero (although their ratio is not constant).[15] Following the remark that both sides of the solutions have a similar shape, it could be that restricting the solution of $v(\cdot)$ to be concave, departing from zero to both sides, could give us an unique solution, if the goal

[14] An easy way to see this is to check for the extreme agents $\alpha = 0$ and $\alpha = 1$.

[15] For the solution depicted in solid line, a guess function $v_o(j) = |j|$ was used. For the other two examples, signified by dashed and pointed lines, more concave curves (always symmetric in zero) were used. The asymmetry is reached by the algorithm itself.

were to find only one solution.

Second, an interesting feature to note is that this difference on both sides of $v(\cdot)$ means that a fully homo-oeconomicus parent has a greater 'dislike' (or disutility) of having a fully Kantian child than the way around. Linked to this observation and to the previous paragraph, we also find that the ratio of the two extremes of $v(\cdot)$, namely $v(-1)/v(1)$, is quite close to the ratio of the two clusters of people, 0.2981/0.3611, no matter which solution of $v(\cdot)$ we use (for this given transformation between γ and α).[16] These two clusters, as was shown in Fig. 2b, agree with the semi-uniform part of the distribution of $f(\alpha)$.

Recalling that having population dynamics with direct and oblique transmission means that a parent will exert effort depending on the actual distribution of the population and the shape of the function $v(\cdot)$, we could think of this function as being a distillation of the parents dislike or disutility of having a child with a different trait compared to theirs. Therefore, this means that the result of having a population with a (slight) majority of homo-oeconomicus people comes, at least in part, from the fact that these people consider being Kantian to be a much worse option. One line of thought that could be explored is the following: it seems that homo-oeconomicus people are selfish compared to Kantian ones, since the latter care for other people. Therefore, it seems reasonable to think that homo-oeconomicus people use a stronger myopic empathy when evaluating their offspring's utility, as compared to Kantian parents. Since the essence of the function $v(\cdot)$ comes from this myopic empathy, this explanation fits well the difference between the positive and negative side of the function $v(\cdot)$: They are the homo-oeconomicus dislike and the Kantian one, respectively. It would be interesting to better understand the social and psychological reasons behind this point, although this vein of thought escapes the scope of the present paper.

Another compelling reason to explore this last point more precisely would relate to understanding how to change a society that is not exhibiting green behaviour into a green one, in line with the topic discussed in Cerda (2015) [3]. If there is a way to modify the function $v(\cdot)$ through education, active information or another social manner, then the society would become more Kantian, which would in turn make it greener.

[16] Using other transformations between γ and α, as in Appendix E will change this ratio, and in that case this statement might not hold true any more. This might signal that the distribution shown in Fig. 2b would be a better fit.

4 Conclusions

The present work connects the trait transmission mechanism to the well-known results in Dictator Games experiments. To do so, first I develop a continuous version of a trait transmission process and I solve for the question: Which trait 'transition function' $v(\cdot)$ would make a population evolve into a specific distribution of that trait? To answer this question, I find the condition that has to be met and develop an algorithm for those cases where analytical solutions are not possible.

Then by mapping the results of the DG into a Kantian trait distribution, I rephrase previous question into the following one: What force is behind this observed distribution of a specific trait, that trait being in this case the the Kantian morale? This force has to do with parents' desire, more or less, of having their offspring resemble themselves. It turns out, as was already clear in the trait transmission literature, that this force shapes the constitution of the society. In the specific case of a Kantian trait and using DG experiments' results, I find that homo-oeconomicus people have a stronger dislike or disutility of having a child with a different trait as compared to their Kantian counterparts. Following the origin of the parent's will behind the trait transmission, we know that myopic empathy is (at least one) reason for wanting our children to be as we are. Hence, homo-oeconomicus people seem to be more myopic than the Kantian ones, which make sense when we consider the definitions of being Kantian and homo-oeconomicus. Homo-oeconomicus people tend to be more selfish, where Kantian ones are more empathic. It turns to be ironic, assuming all these assumptions as true, that Kantian people, in caring more for their fellows, are jeopardizing their own (evolutionary) existence.

As mentioned in the previous section, it would be interesting to better understand the origins of the function $v(\cdot)$, the one responsible for the transmission forces that shape our society, at least in the Kantian arena. A better understanding is not only appealing for its own merits, but could also help to figure out how to move societies to become greener. A second line of future research could to find out a more general mathematical solution of the function $v(\cdot)$.

References

[1] Alberto Bisin and Thierry Verdier. A model of cultural transmission, voting and political ideology. *European Journal of Political Economy*, 16(1):5–29, March 2000.

[2] Alberto Bisin and Thierry Verdier. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2):298 – 319, 2001.

[3] Lorenzo Cerda Planas. Moving toward greener societies: Moral motivation and green behaviour. Working paper, Paris School of Economics - Université Paris 1 Panthéon - Sorbonne, March 2015.

[4] Christoph Engel. Dictator games: a meta study. *Experimental Economics*, 14(4):583–610, 2011.

[5] Daniel Friedman. On economic applications of evolutionary game theory. *Journal of Evolutionary Economics*, 8(1):15–43, 1998.

[6] Marc Harper. Information geometry and evolutionary game theory. *arXiv preprint arXiv:0911.1383*, 2009.

[7] Esther Hauk and Maria Saez-Marti. On the cultural transmission of corruption. *Journal of Economic Theory*, 107(2):311 – 335, 2002.

[8] Josef Hofbauer and Karl Sigmund. Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 40(4):479–519, 2003.

[9] Julie Novakova and Jaroslav Flegr. How much is our fairness worth? the effect of raising stakes on offers by proposers and minimum acceptable offers in dictator and ultimatum games. *PLoS ONE*, 8(4), 2013.

[10] Tanya O'Garra and Dave Krantz. Bystander effects and compassion fade: An experimental investigation of how giving varies with the number of dictators and recipients. *Working paper presented at 5th WCERE Conference (World Congress of Environmental and Resource Economics), Istanbul, Turkey*, 28 June-2 July 2014.

[11] Maria Sáez-Martı and Yves Zenou. Cultural transmission and discrimination. *Journal of Urban Economics*, 72(2):137–146, 2012.

[12] Karl Sigmund. A survey of replicator equations. In *Complexity, Language, and Life: Mathematical Approaches*, pages 88–104. Springer, 1986.

[13] Gerald Silverberg. *Evolutionary modeling in economics: recent history and immediate prospects*. Citeseer, 1997.

[14] E Christopher Zeeman. Population dynamics from game theory. In *Global theory of dynamical systems*, pages 471–497. Springer, 1980.

A General evolution equation for a n type population

Having an n-type population, let us call q_t^i the share of type i at time t. Recalling the definition of the transition probabilities given in Eqn. (2.5), we have:

$$P^{ii} = \tau^i + (1-\tau^i)q_t^i \qquad P^{ij} = (1-\tau^i)q_t^j \qquad \forall i \neq j \tag{A.1}$$

Therefore, the share of type i at time $t+1$ will simply be:

$$\begin{aligned}
q_{t+1}^i &= P^{ii}q_t^i + \sum_{j\neq i} P^{ji}q_t^j \\
&= \Big(\tau^i + (1-\tau^i)q_t^i\Big)q_t^i + \sum_{j\neq i}\Big((1-\tau^j)q_t^i\Big)q_t^j \\
&= q_t^i\Big(\tau^i + (1-\tau^i)q_t^i + \sum_{j\neq i}(1-\tau^j)q_t^j\Big) \\
&= q_t^i\Big(\tau^i + \sum_{j=1}^{n}(1-\tau^j)q_t^j\Big) \\
&= q_t^i\Big(\tau^i + \sum_{j=1}^{n} q_t^j - \sum_{j=1}^{n}\tau^j q_t^j\Big) \\
q_{t+1}^i &= q_t^i\Big(\tau^i + 1 - \overline{\tau}\Big)
\end{aligned}$$

$$\Delta q_t^i = q_{t+1}^i - q_t^i = q_t^i\,(\tau^i - \overline{\tau}) \qquad\qquad \text{with} \qquad \overline{\tau} = \sum_{i=1}^{n}\tau^i q_t^i$$

B Solving for an specific $f(\alpha)$

I find some solutions $v(j)$ that solve problem defined by the Eqn. (2.11), rewritten here:

$$\int_0^1 v(j,\alpha)f(\alpha)d\alpha = \tau(j) = \text{constant} \qquad \forall j \mid f(j) > 0 \tag{B.1}$$

for $f(\alpha) = C_1\,\delta(0) + C_2 + C_1\,\delta(1)$ with $2C_1 + C_2 = 1$. One way to solve this integral equation is to make use of non-standard calculus and performing the computations within the *hyperreals* (approach that is more straightforward). With this set-up, we have that $f(0)$ (where the first Dirac delta is centred) is equal to a infinite hyperreal, such that $f(0)\cdot\epsilon = C_1$, with ϵ being an infinitesimal hyperreal. In the same fashion, we have that $f(1) \cdot \epsilon = C_2$. Hence,

$$\begin{aligned}
\tau(j) = \int_0^1 v(j,\alpha) f(\alpha) d\alpha &= v(j)\, f(0)\, \epsilon \,+\, C_2 \int_\epsilon^j v(j-\alpha) d\alpha + C_2 \int_j^{1-\epsilon} v(j-\alpha) d\alpha \,+\, v(1-j)\, f(1)\, \epsilon \\
&= v(j)\; C_1 + C_2 \Big(\int_\epsilon^j v(j-\alpha) d\alpha + \int_j^{1-\epsilon} v(j-\alpha) d\alpha \Big) + v(1-j)\; C_1 \\
\text{(returning to the real domain)} &= C_1 \big(v(j) + v(1-j) \big) + C_2 \Big(\int_0^j v(j-\alpha) d\alpha + \int_j^1 v(j-\alpha) d\alpha \Big) \\
\text{(changing variable inside integrals)} &= C_1 \big(v(j) + v(1-j) \big) + C_2 \Big(\int_0^j v(x) dx + \int_0^{1-j} v(x) dx \Big) \\
\text{(replacing } w(j) = \int_0^j v(x) dx\text{)} &= C_1 \big(v(j) + v(1-j) \big) + C_2 \big(w(j) + w(1-j) \big) \qquad \text{(B.2)}
\end{aligned}$$

which has to be constant (equilibrium condition).

Finding general solutions for this problem can be a colossal task. One approach to find some solutions is to look for them in following equations:

$$C_1\, v(j) + C_2\, w(j) = \text{constant} \qquad \text{for } 0 \leq j \leq 1 \qquad \text{(B.3)}$$

$$C_1\, v(j) + C_2\, w(1-j) = \text{constant} \qquad \text{for } 0 \leq j \leq 1 \qquad \text{(B.4)}$$

It is easy to see that solutions for (B.3) and (B.4) are also solutions for (B.2). This approach obviously restricts the possible solutions to be found, but it also eases that task at hand.

Using (B.3), I found that the family of functions $v(j) = K - a \cdot c \cdot e^{-a|j|}$ (with $a = C_2/C_1$) is a solution for $f(\alpha) = C_1\, \delta(0) + C_2 + C_1\, \delta(1)$ with $2C_1 + C_2 = 1$. I verify this by checking that the following expression is constant:

$$\begin{aligned}
\tau(j) &= C_1 \big(v(j) + v(1-j) \big) + C_2 \big(w(j) + w(1-j) \big) \\
&= C_1 \big(v(j) + v(1-j) \big) + C_2 \Big(\int_0^j v(x) dx + \int_0^{1-j} v(x) dx \Big) \\
&= C_1 \big(K - ac\, e^{-a|j|} + K - ac\, e^{-a|1-j|} \big) + C_2 \Big(\int_0^j K - ac \cdot e^{-a|x|} dx + \int_0^{1-j} K - ac \cdot e^{-a|x|} dx \Big) \\
&= 2KC_1 - acC_1 e^{-aj} - acC_1 e^{-a(1-j)} + C_2 \Big(Kx \Big|_0^j + ce^{-ax} \Big|_0^j + Kx \Big|_0^{1-j} + ce^{-ax} \Big|_0^{1-j} \Big) \\
&= 2KC_1 - acC_1 e^{-aj} - acC_1 e^{-a(1-j)} + C_2 \Big(Kj + ce^{-aj} - c + K(1-j) + ce^{-a(1-j)} - c \Big) \\
&= 2KC_1 + C_2 K - 2C_2 c + ce^{-aj}(C_2 - C_1 a) + e^{-a(1-j)}(C_2 - C_1 a) \\
&= 2KC_1 + C_2 K - 2C_2 c \qquad \text{(since } a = C_2/C_1\text{)} \\
&= \text{constant}
\end{aligned}$$

In a similar way, it can be shown that $v(j) = b \cdot sin\left(\frac{(4k-3)\pi}{2} \cdot |j|\right)$ is a solution. In this case, we focus on solving for (B.4). Because of the properties of the trigonometric functions, in particular their derivatives and reflections in $\pi/2$,[17] we are able to find these types of solutions. The idea is to make the range of $[0, 1]$ of j coincide, with a change of variable, with $[0, \pi/2]$, $[0, 5\pi/2]$, etc. The proof follows the same line as the one before and it is left to the reader.

C Properties of the Algorithm

As stated in Section 3, I show that if at some iteration k, $v_k(j)$ is a solution, the algorithm converges, and that if it converges then $v_k(j)$ is a solution.

Let $v_k(j)$ be a solution. This means that $\tau_k(j)$ is constant for $0 \leq j \leq 1$. Therefore, the term $1/\tau_k(j)$ is also constant, and hence the adjustment function $a_k(j) = \frac{1}{\tau_k(1-2j)}$ is also constant. Let call this last constant K. From the iteration process we see that:

$$v_{k+1}(j) = a_k(j) \cdot v_k(j) = K \cdot v_k(j)$$

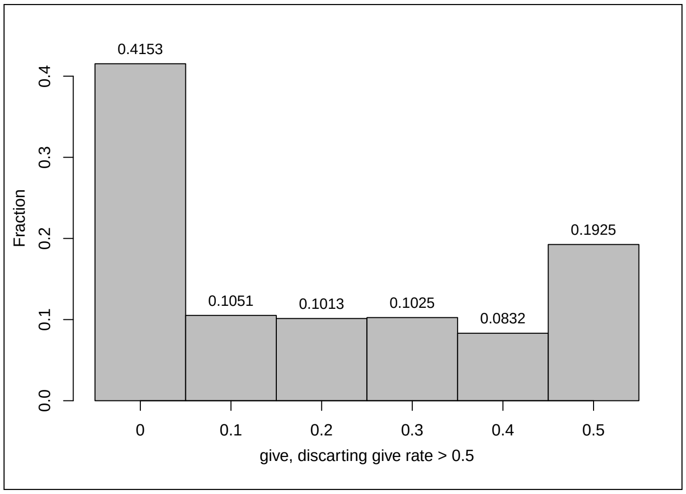
It is easy to see that $K = 1$. If not, having $K < 1$, would yield to $v_k(j) = 0$, when $k \to \infty$, which would mean $\tau_\infty(j) = 0$ and then $a_\infty(j) = \infty$, which is a contradiction. Similarly, if $K > 1$ we get that $a_\infty(j) = 0$, again a contradiction. Therefore, $K = 1$.

On the other hand, if the iteration converges, it means that $v_{k+1}(j) = v_k(j)$ for k bigger than some fixed number. This means that $a_k(j) = 1$, which in turn means that $\tau_k(j) =$ constant, proving that $v_k(j)$ is a solution.

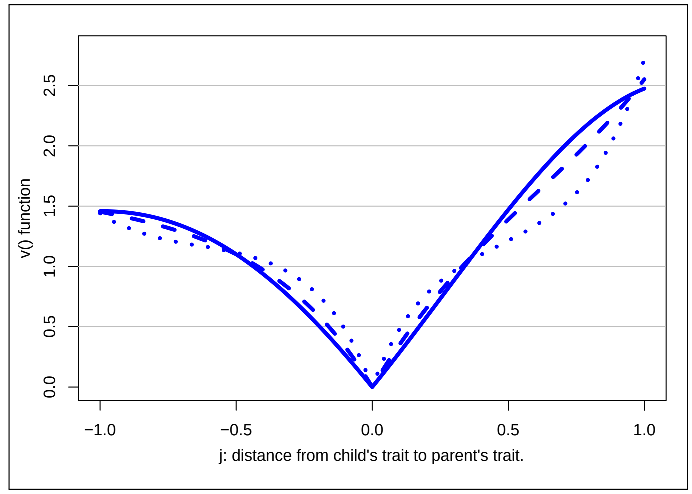
D Alternative Case: Disregarding give rates over 0.5

Here I find the $v(\cdot)$ solution for the Kantian trait of the DG, using a truncated version of Engel's (2011) [4] data. As stated in Section 3, here I perform the same operations, but I discard all give rates bigger than 50%. To do this, I drop this information and recalculate the fractions of give rates within this new universe. I then transform this information into its Kantian distribution equivalent, and finally I find the solution function $v(\cdot)$ using the algorithm previously mentioned. The histogram of the truncated information (give rate) and some solutions are plotted in the following figure. To produce these solutions, the same initial guesses for $v_0(\cdot)$ were used as in Section 3. As we can observe, the solution $v(\cdot)$ is once again asymmetrical, although its 'sides' are yet alike. The difference between

[17] Meaning, for example, that $sin(\pi/2 - \theta) = cos(\theta)$.



(a) Truncated distribution of give rates, using Engel (2011).



(b) Solutions of $v(\cdot)$ for the corresponding Kantian distribution.

Figure 4: Truncated data for give rates and some solutions of $v(\cdot)$.

this and the function depicted in Fig. 3 (Page 13) is that in this one, the two sides ($v(+)$ and $v(-)$) are more different, following the stronger dissimilarity of the new clusters in $\alpha = 0$ and $\alpha = 1$ (Fig. 4a). However, the essential result holds: the homo-oeconomicus agent has a stronger dislike of having a Kantian child than the opposite case.

E Using different utility functions for transforming give rate into Kantian trait.

As stated in Section 3, choosing different utility functions in Eqn. (3.6) produces different transformations from give rate γ into Kantian trait α. Always using a CRRA function, I depict different transforms for different values of ϵ, their risk aversion measure:

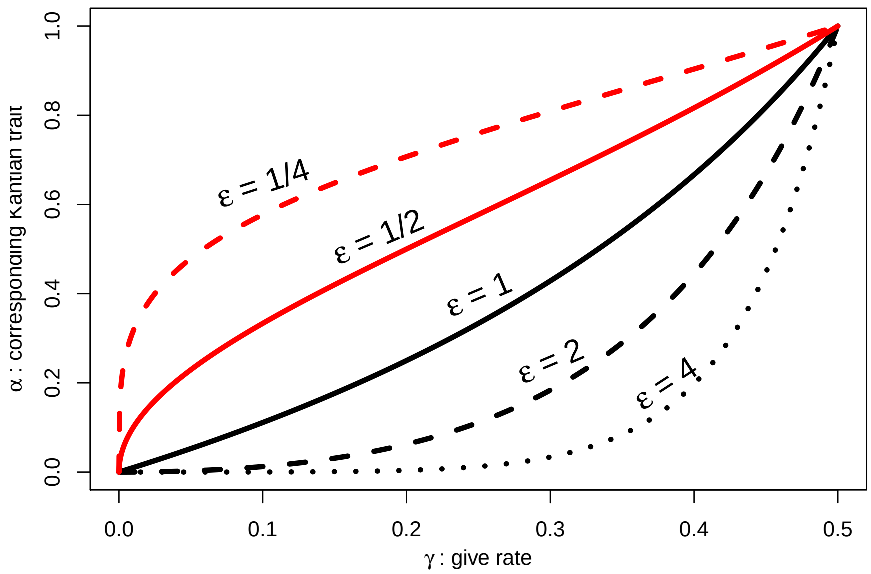
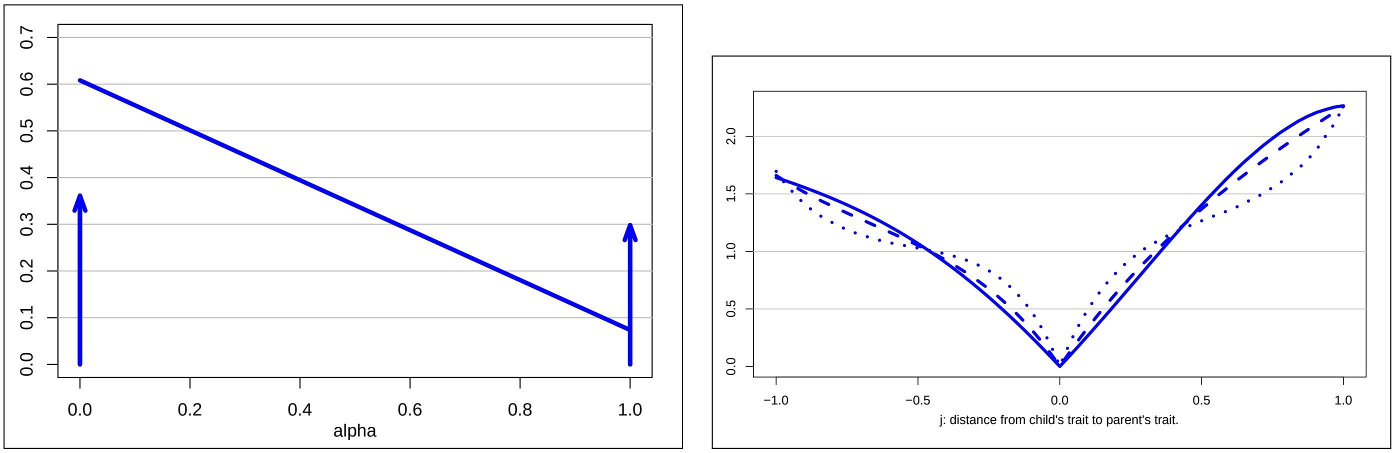


Figure 5: Examples of transforms from give rate to Kantian trait.

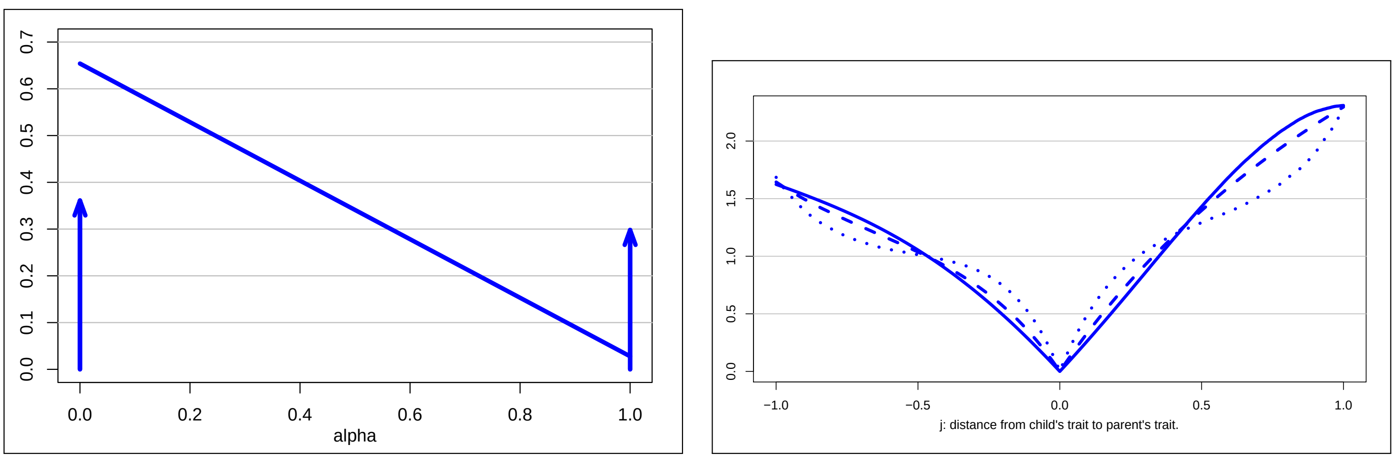
As one can observe from Fig. 5, using measures of risk aversion too different from $\epsilon = 1$ produces a loss of information. For example, with $\epsilon = 4$, values of give rate between zero and 0.25 yield to almost the same Kantian trait, around zero. We find a similar result if we use bigger values, as in the case of $\epsilon = 1/4$.

I turn now into how the distribution of the Kantian trait α could be, depending on the choice of ϵ and then, how this would translate info the solution of $v(\cdot)$. In the following figures I plot the case with $\epsilon = 0.9$ and $\epsilon = 1$:



(a) Distribution in Kantian measurement with $\epsilon = 0.9$ (b) Solution $v(\cdot)$, with $\epsilon = 0.9$

Figure 6: Results using $\epsilon = 0.9$



(a) Distribution in Kantian measurement with $\epsilon = 1$ (b) Solution $v(\cdot)$, with $\epsilon = 1$

Figure 7: Results using $\epsilon = 1$